HHS Public Access

Author manuscript

Neurogastroenterol Motil. Author manuscript; available in PMC 2016 August 01.

Published in final edited form as:

Neurogastroenterol Motil. 2015 August ; 27(8): 1110–1119. doi:10.1111/nmo.12590.

Agreement between Prospective Diary Data and Retrospective Questionnaire Report of Abdominal Pain and Stooling Symptoms in Children with Irritable Bowel Syndrome

Mariella M. Self, Ph.D.[1,2,3], **Amy E. Williams, Ph.D.**[4], **Danita I. Czyzewski, Ph.D.**[1,2,3], **Erica M. Weidler, M.Ed.**[1,3,5], and **Robert J. Shulman, M.D.**[1,3,5]

[1]Department of Pediatrics, Houston, Texas

[2]Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, Texas

[3]Texas Children's Hospital, Houston, Texas

[4]Department of Psychiatry, Indiana University School of Medicine; Psychiatry & Behavioral Sciences, Riley Hospital for Children, Houston, Texas

[5]Children's Nutrition Research Center, Houston, Texas

Abstract

Background—In functional gastrointestinal disorders, patient recall of symptoms drives diagnostic decisions and evaluation of treatment response, as well as research conclusions about potential treatments. In pediatrics, parent report also impacts assessment and care. Hence, identifying methods for accurately capturing patient and parent report of Irritable Bowel Syndrome (IBS) symptoms is important. This study evaluated correspondence between retrospective questionnaire (parent and child report) and prospective diary data for children and adolescents with IBS.

Methods—Participants included 50 children/adolescents with IBS per Rome III criteria. Children completed a two-week pain and stool diary. Children and parents subsequently completed a two-week recall questionnaire, reporting number of pain days, maximum pain, days without bowel movement, and days with diarrhea during the diary interval. Intraclass correlation coefficients and Bland-Altman plots assessed agreement.

Key Results—For pain and days without bowel movement, overall agreement between child recall questionnaire and child diary was strong, though under conditions likely to facilitate

Corresponding Author: Mariella M. Self, Ph.D., Texas Children's Hospital, Clinical Care Center, Ste.1740, 6701 Fannin St. CC1740.01, Houston, TX 77030-2399, mmself@texaschildrens.org, (832) 822-3752 phone, (832) 825-3747 fax.

DISCLOSURES

Competing Interests: the authors have no competing interests.

Specific author contributions:

MS – study concept and design; statistical analysis and interpretation of data; drafting of the manuscript
AW – statistical analysis and interpretation of data; critical revision of the manuscript for important intellectual content
DC - interpretation of data; critical revision of the manuscript for important intellectual content
EW - acquisition of data; critical revision of the manuscript for important intellectual content
RS - obtained funding; study concept; interpretation of data; critical revision of the manuscript for important intellectual content

agreement and with individual variation observed. Parent recall and child diary were less concordant, and agreement about diarrhea was poor for parent and child. Age did not significantly correlate with agreement.

Conclusions & Inferences—Child questionnaire with short recall interval may be a reasonable approximation for diary data, though this varies by individual and replication/investigation of lengthier recall are needed. Relying on parent questionnaire does not appear a suitable proxy, and recall of stool form by both parent and child appears more problematic. These results combined with existing literature support use of diary data whenever possible.

Keywords

functional gastrointestinal disorders; irritable bowel syndrome; Rome III questionnaire; pain diary; children and adolescents

Clinical decision-making about diagnosis and treatment is routinely guided by patient report of symptoms, with symptom recall critically influencing medical management particularly for conditions lacking objective indicators such as functional gastrointestinal disorders (FGIDs). In research, patient-reported outcomes are increasingly conceptualized as primary endpoints[1, 2], emphasizing patient report in evaluating potential treatments.

Health-related information recalled via questionnaire has been questioned as unreliable or subject to bias.[3] For example, peak and end effects (i.e., most intense and most recent pain) disproportionately influence pain recall and can undermine validity of retrospective self-reports.[3–13] Recalled pain ratings are typically higher than momentary assessments, and lengthening recall interval magnifies bias.[4,9, 11,13–19] Despite these limitations, retrospective questionnaires are used routinely and endorsed for assessing symptoms in FGID treatment trials[1]. Given this endorsement and the burden of diaries[20], evaluating if questionnaires are a sufficient proxy is worthwhile.

Most literature evaluating correspondence between recalled and recorded symptoms involves non-GI samples, but evidence suggests GI symptoms are subject to recall error, or that recalled and recorded bowel function diverge.[21–27] This small literature is often limited by lack of correspondence between questionnaire and diary intervals (e.g., questionnaires preceding diary). This challenges interpretation, but adult IBS literature suggests that subtyping differs based on recalled versus recorded data[26,27]. Rome III Diagnostic Questionnaires may overestimate the frequency of abnormal stool form[26], participants tend to recall more extreme stool forms as representative[27,28], and those describing constipation underestimate stool frequency on questionnaire.[23,24] Lackner and colleagues recently reported that, though as a group adults with IBS accurately recalled some IBS symptoms, individual correspondence varied with a subset of patients evidencing poor recall accuracy.[28]

Few studies examine correspondence between recalled and recorded pain in children, particularly in GI samples. Available research employs variable methods and yields variable interpretations concerning children's recall accuracy. Accuracy usually increases with age, recalled pain ratings tend to be higher than momentary ratings (though pediatric findings are

more mixed), and peak- and end-effects similarly bias ratings.[14,18,29–32] Even less pediatric research concerns recalled versus recorded stool data, though evidence suggests recalled and recorded defecation frequency do not closely correspond and methodology affects diagnostic classification.[33]

One study by Chogle and colleagues[34] examined correspondence between recalled and recorded pain in pediatric FGIDs, comparing a four-week pain diary to retrospective report of number of pain days. Results reflected a moderate positive correlation (Spearman correlation = 0.4), with 16% of children having perfect agreement. Interestingly, younger children had higher correspondence than adolescents. Group data reflected more pain days via diary than recall, but individual data indicated that 54% of children recalled fewer episodes than reported on diary, whereas 40% recalled more.[34]

To extend scant and often methodologically limited literature, we evaluated correspondence between retrospective questionnaire and prospective diary in pediatric IBS, evaluating pain and stool variables during the same reporting interval. Given the importance of parent perception in pediatrics and evidence that parent- and child-report differ[35], correspondence between parent questionnaire and child diary was also examined. Relationship of child age to correspondence between child diary and questionnaire was also assessed.

MATERIALS & METHODS

Recruitment & Participants

Participants for this study included 50 children and adolescents ages 8–18 years with IBS as defined by Pediatric Rome III criteria who were participating in a double-blind randomized controlled six-week trial of soluble fiber versus placebo. Participants for the study were recruited from primary (n = 22) and tertiary care (n = 28) clinics in a large academically-affiliated pediatric health care network. Potential participants were identified through billing records documenting ICD-9 codes for abdominal pain or IBS. Medical charts were reviewed and potentially eligible participants were invited to participate by their physician via letter and, if interested in participating, were screened by telephone for full eligibility. Eligibility required that parent and/or child (if child ≥11 years) responses met criteria for IBS on a telephone-administered screening questionnaire based on Rome III criteria.

Children were excluded from participation if chart review or telephone screening revealed that symptoms were accounted for by organic disease (or organic disease remained in the differential), a significant chronic health condition requiring daily medication or specialty follow up care, vomiting ≥ 2 times per month within the preceding three months, or unintentional weight loss of ≥5% body weight within a 3-month period. Additional exclusion criteria included lack of fluency in English (as the parent study required completion of questionnaires only available in English), cognitive impairment significantly below age and/or grade level, significant mental health diagnoses (e.g., Bipolar disorder), or current participation in psychotherapy for abdominal pain.

Measures

Two-Week Pain & Stool Diary—Using a paper diary, children rated abdominal pain for three intervals each day (morning, midday/afternoon, and evening/nighttime) using a 0–10 numerical scale anchored with the phrases “no pain at all” and “the worst pain you can imagine.” Children recorded the time of each stool and rated its consistency using the Bristol Stool Form Scale (BSFS).[35] Stool form ratings of 6 or 7 were classified as representing occurrences of diarrhea. Children called their responses into a dedicated phone line linked to a computerized database at the end of each day.

For the purposes of this study, diary variables included each participant’s number of pain days (i.e., days where pain was rated as ≥ 1 on the 0–10 scale for any of the 3 rating intervals), maximum pain rating, number of days with no bowel movement recorded, and number of days with diarrhea recorded during the two week diary interval.

Two-Week Recall Questionnaire—After completion of the prospective diary, children and parents used a two-week recall interval commensurate with the diary to indicate how many days stomach pain or discomfort occurred (0–14 days) and to indicate the maximum abdominal pain experienced during the two-week diary period using the 0–10 numerical scale. Children and parents were also asked how many days (0–14) the child had no bowel movement and on how many days the child had diarrhea during the two-week diary interval.

Procedure

The study was approved by the Baylor College of Medicine Institutional Review Board, and parent consent and child assent were obtained. Though parents and children consented/assented to complete a variety of questionnaires and were aware the questionnaires would be repeated at study end, they were not explicitly cued that they would be asked to recall pain and stool information. During a home visit, parents and children were instructed on completion of the two-week daily pain and stool diary. Parents were asked to remind children to complete the diaries daily but were instructed to allow children to independently rate abdominal pain and record stool occurrence and form without influencing child responses. Data for this study utilized post-treatment diaries that were completed during the last two weeks of the six-week treatment phase. Because this study concerned correspondence between observed and recalled pain and stool data, treatment assignment would not be expected to differentially affect correspondence.

Following completion of the two-week diary, parents and children independently responded to the two-week recall questionnaire administered via telephone by a trained research assistant. Parents and children were explicitly cued that recall pertained to the two-week diary interval and were asked to separate when providing responses so as to prevent influencing each other’s report. Anecdotally this appeared to be complied with, as a time delay or calling out for the other reporter routinely preceded switching the phone to the other reporter. Parents and children were instructed not to reference the diary when completing the recall questionnaire and at the end of the questionnaire were asked if they had used the diary to answer any questions (all denied doing so). Participants were excluded from data analysis

if the interval between diary completion and administration of the recall questionnaire exceeded 6 days.

Statistical Analysis

Dependent-samples t-tests compared child diary report, child questionnaire report, and parent questionnaire report for the pain and stooling variables. To facilitate comparison with the Chogle et al. study[34], Spearman's rank order correlation coefficients were computed, and scatter plots were also visualized. However, because correlation coefficients are based on ordering and relative distances between intervals, they do not account for the possibility of differences in variability or mean level between measures (e.g., if pain recall were higher on questionnaire vs. diary[17]). Intraclass correlation coefficients were developed to compare two measures intended to measure the same construct using the same scale by taking variability and mean level into account. Therefore, single measures intraclass correlation coefficients (ICCs; two-way mixed models with absolute agreement) were used to assess agreement between retrospective questionnaire recall and diary observations for four variables: number of pain days, maximum pain rating, number of days with no bowel movement, and number of days with diarrhea during the two-week diary interval. Agreement was assessed for both parent and child recall, as compared with child-reported diary observations. Guidelines for interpreting the clinical significance of ICCs have been recommended as follows: <0.40=Poor; 0.40–0.59=Fair; 0.60–0.74=Good; and ≥0.75=Excellent.[37,38] Difference scores between diary and questionnaire variables were computed to evaluate the direction of discrepancies (i.e., if diary or questionnaire yielded higher symptom ratings), with the percentage of the sample evidencing equivalent ratings and each discrepancy reported. To evaluate if age was related to agreement for child questionnaire versus child diary, Pearson correlation coefficients were computed for age and the difference scores (i.e., questionnaire minus diary) for the four study variables. Bland-Altman plots were also constructed and visually inspected to assess the degree of correspondence between the recall questionnaire and prospective diary assessment methods.[39]

RESULTS

Demographics

Twenty-four participants were excluded from data analysis because the interval between diary completion and administration of the recall questionnaire exceeded 6 days. The remaining 50 study participants (fiber = 28; placebo = 22) had a mean age of 13.4 ± 2.6 years and were 62% female. Distribution of race/ethnicity was 68% White, 24% Hispanic, and 6% Black.

Descriptive Statistics

Means and standard deviations for child diary report, child questionnaire report, and parent questionnaire report for the pain and stooling variables are provided in Table 1. Dependent samples *t*-tests reflected that as a group, parents reported less pronounced pain symptoms via retrospective questionnaire than were reported by children via either diary or questionnaire. For stooling symptoms, significantly more days without a bowel movement

were recorded on children's diaries than reported by either child- or parent-report retrospective questionnaire. No difference in number of days with diarrhea emerged between reporters or methods, and overall, diarrhea was an infrequently recorded occurrence on the two-week diary (thirteen participants, seven of whom reported only 1 day of occurrence).

Agreement for Number of Pain Days

The Spearman's rank order correlation coefficient for number of pain days based on diary report versus child-report questionnaire was .82 and the ICC was .83 (i.e., "excellent"), with 24% of children evidencing exact agreement between recalled and recorded number of pain days (Table 2). Discrepancy occurred in both directions, with diary report of pain days exceeding questionnaire recall being slightly more common. The simple scatter plot and the Bland-Altman plot evaluating child recall versus diary for number of pain days are provided in Supplemental Figure 1. The Bland-Altman plot did not reflect systematic method bias (i.e., one method of assessment was not consistently higher than the other) but demonstrated that the upper and lower limits of agreement encompassed a fairly wide range of values with four data points that exceeded those limits. The plot reflected higher agreement at lower frequency of pain.

For parent recall of number of pain days, the Spearman's rank order correlation was .53 and the ICC was .55 (i.e., "fair"), also with 24% of parents having exact agreement (Table 2). The most common direction of discrepancy was diary report of pain days exceeding questionnaire recall. The simple scatter plot and the Bland-Altman plot comparing parent recall and child diary report of pain days are provided in Supplemental Figure 2. The Bland-Altman plot again did not reflect systematic bias between methods but demonstrated a wide range of values within the upper and lower limits of agreement, with 4 data points that exceeded those limits, and agreement was higher at lower frequency of pain.

Agreement for Maximum Pain Rating

The Spearman's rank order correlation coefficient for maximum pain rating based on child-reported diary versus child-report questionnaire was .76 and the ICC was .80 (i.e., "excellent"), with 40% of children having exact agreement (Table 2). Discrepancy occurred fairly equally in both directions. The simple scatter plot and Bland-Altman plot are provided in Supplemental Figure 3. The Bland-Altman plot indicated no systematic method bias and that agreement was not notably different across levels of pain. Two data points exceeded the upper and lower limits of agreement.

Comparing child diary and parent recall for maximum pain, the Spearman's rank order correlation coefficient was .42 and the ICC was also .42 (i.e., toward the lower limit of "fair" agreement), with 22% of parents having exact agreement with child diary (Table 2). Discrepancy again occurred fairly equally in both directions. The simple scatter plot and Bland Altman plot for parent recall of maximum pain rating are provided in Supplemental Figure 4. The Bland-Altman plot again indicated no systematic method bias and that agreement was not notably different across levels of pain. The upper and lower limits of agreement contained a broader range of values for parent compared with child questionnaire report. One data point exceeded the upper and lower limits of agreement.

Agreement for Number of Days without a Bowel Movement

When comparing the number of days on which no bowel movement occurred, the Spearman's rank order correlation coefficient for child diary versus child-report questionnaire was .77 and the ICC was .74 (i.e., "good"). Twenty-four percent of children evidenced exact agreement, and though both directions of discrepancy occurred, diary report of more days without stool than were recalled via questionnaire was notably the more common discrepancy (Table 2). The simple scatter plot and Bland-Altman plot are provided in Supplemental Figure 5. The Bland-Altman plot did not reflect systematic method bias but demonstrated that the upper and lower limits of agreement encompassed a fairly wide range of values. One data point exceeded those limits. Agreement tended to be higher when days with no bowel movements occurred less frequently.

The Spearman's rank order correlation coefficient for child diary versus parent recall was .57 and the ICC was .74 (i.e., "good"). Sixteen percent of parents exactly recalled the number of days their child experienced no bowel movement in the two-week interval. Similar to the child-report results, the most common direction of discrepancy was more days with no stool on diary than recall questionnaire (Table 2). The simple scatter plot and Bland-Altman plot are provided in Supplemental Figure 6. Systematic method bias again was not found, but the upper and lower limits of agreement again encompassed a fairly wide range of values with two data points that exceeded those limits. Agreement tended to be higher when number of days with no bowel movement was lower.

Agreement for Number of Days with Diarrhea

The Spearman's rank order correlation coefficient for child diary versus child recall of number of days with diarrhea compared with diary observation was .12, and the ICC was −.03 (i.e., "poor"). However, the methods yielded exact agreement for 58% of children, with questionnaire recall exceeding diary for report of diarrhea more often than the reverse (Table 2). As stated previously, diary recording of diarrhea was a relatively infrequent occurrence (i.e., 13 participants); all but one of those evidencing exact agreement between diary and questionnaire were recalling absence of diarrhea. The simple scatter plot and Bland-Altman plot are provided in Supplemental Figure 7 and demonstrate that a few significant outliers in both directions greatly affected the results. Systematic method bias did not occur but the upper and lower limits of agreement again encompassed a fairly wide range of values with four data points that exceeded those limits. Agreement was markedly better when diarrhea was not present.

For parent recall, the Spearman's rank order correlation coefficient was .16 and the ICC was .07 (i.e., "poor"). However, 58% of parents evidenced exact agreement, and both directions of discrepancy were reasonably equally represented (Table 2). The simple scatter plot and Bland-Altman plot are provided in Supplemental Figure 8, again demonstrating that a few significant outliers in both directions notably impacted the results. Systematic method bias did not occur but the upper and lower limits of agreement again encompassed a fairly wide range of values with three data points exceeding those limits. Agreement again was markedly better when diarrhea was not present.

Relationship of Age with Agreement

Pearson correlations coefficients indicated that age was not significantly related to the difference scores (i.e., child questionnaire minus child diary) for any pain or stool variable (Table 3).

Distribution of Discrepancy for Child-Reported Pain and Stooling Variables

To evaluate if a subset of consistently accurate recall reporters could be identified or if children tended to evidence discrepancy in a consistent manner (e.g., questionnaire exceeding diary for multiple variables), we evaluated the distribution of discrepancy for pain and stooling variables (Table 4). Overall, results for child report indicated that discrepancies did not reflect a consistent response set. For example, for 12% of the sample questionnaire exceeded diary report for both number of pain days and maximum pain rating, whereas for 10% of the sample diary exceeded questionnaire for number of pain days but questionnaire exceeded diary for maximum pain rating. Only one child provided equivalent report via both mechanisms for all variables. For only two children did diary exceed questionnaire report for all four variables, and for only one child did questionnaire exceed diary report for all variables.

DISCUSSION

Given that recall of symptoms and description of their change over time drives diagnostic decisions and conclusions about treatments in FGIDs, identifying methods that both accurately and feasibly capture patient and parent report of symptoms is important. Our study extends the scant literature comparing recalled and recorded data in FGIDs with methodological improvements including correspondence between questionnaire and diary reporting interval, use of ICCs, inclusion of stool variables in addition to pain report, and evaluation of both parent and child recall.

When examining group data, we found that under our study conditions, child retrospective two-week questionnaire recall of pain and days without a bowel movement were reasonably concordant (Table 2). However, visual inspection of the scatter plots and Bland-Altman plots demonstrate considerable individual variability in correspondence between the two assessment methods (Table 2; Supplemental Figures 1, 3, 5). Such variability in accuracy is particularly important when considering clinical decision-making at the individual patient level, as some patients' recall will not correspond well to observational data, and discriminating which patients are accurate reporters is challenging. Further, when interpreting these results, it is important to note that these data were collected under conditions likely to facilitate agreement. For child report in particular, the indices of agreement represent recall under optimal conditions because children were attending to and recording these variables on the diary, again entered their ratings by phone, and then were specifically cued to recall the diary interval and to report symptoms using the same metric (i.e., 0–10 pain scale). The process of monitoring and recording symptoms via the diary may have improved the recall of symptoms. Therefore, these results likely represent higher concordance than would occur under routine conditions, where child-reported recall of symptoms seems likely to be more fallible.

Parent questionnaire recall generally did not correspond very well with child diary data (Table 2). That agreement between methods would decrease when reporters also differ is intuitive, and the subjectivity of pain symptoms and potential lack of parent observation/awareness of child stools also would be expected to decrease agreement. However, parent-report of child symptoms is typically a weighty consideration in pediatric care and may also be the basis of study recruitment. Our data suggest that this may be problematic in children and adolescents with IBS. Our study did not incorporate parent diaries, which may be a helpful future addition to discern discrepancy arising from method versus reporter.

For both parent and child report across variables, our results indicated that discrepancy between recalled and recorded symptoms occurred in both directions. This suggests individual differences in reporting accuracy rather than method bias (Supplemental Figures 1–8) and again emphasizes the individual variability in both parent and children's ability to accurately report pain and stool variables via retrospective questionnaire. Similar to Chogle et al.'s study[34], for child-reported pain days, diary exceeding questionnaire report was the most common direction of discrepancy. Interestingly, age was not significantly related to discrepancy between child questionnaire recall and diary. This is contrary to general indication that symptom recall accuracy increases with age in non-GI samples and also differs from Chogle et al.'s[34] finding that younger children had significantly higher correspondence between recalled and recorded pain than adolescents. Given discrepant results, further evaluation of recalled versus recorded IBS symptoms across developmental level are indicated.

As is observable in Supplemental Figures 7 and 8, the agreement indices for number of days with diarrhea were significantly affected by a few outliers for both child and parent report. Only 26% of children and adolescents in our study reported experiencing diarrhea during the two-week diary interval, with the number of days with diarrhea ranging from 1 (most typically) to 9 days. Though the ICCs were extremely low, the majority of parents and children demonstrated exact agreement between recalled and recorded diarrhea data, most of whom were recalling its absence. Outliers notwithstanding, it is logical that it is easier to accurately recall complete symptom absence than its degree of occurrence.

Our study utilized a relatively short recall interval compared with routine clinical practice, and particularly as compared with the Rome III questionnaire, which inquires about the past two months. Therefore, the degree of correspondence between recalled and recorded symptoms reported here, while already imperfect, is likely to be significantly greater than would occur under less standardized conditions or with a lengthier recall period. Indeed, though our study yielded a relatively similar percentage of children with exact agreement for pain days as was found in the Chogle et al. study[34] (24% vs. 16%, respectively), our results reflected notably higher correspondence between child recall of number of pain days and diary data (i.e., Spearman's rho of .82 vs. .40, respectively), and our shorter recall interval (i.e., 2 rather than 4 weeks) may account in part for this difference. Prior to confidently adopting child recall questionnaires in lieu of diary data, future research should seek to replicate these results and to investigate correspondence between recalled and recorded data under lengthier recall intervals (e.g., the Rome III questionnaire).

Pediatric IBS diagnosis based on the Rome III criteria is predicated on recall of at least two months[40], as well as report of symptom interrelationships that must be present weekly for those two months, both of which likely present challenges in capturing accurate information to inform diagnosis. The FDA's guidance document for patient reported outcomes recommends the use of diaries and either short recall periods or report on current state, explicitly cautioning that asking patients to recall over a long interval likely undermines validity.[41] However, in a condition such as IBS that often presents with inherent symptom variability and fluctuation, a short recall interval could underestimate or misrepresent symptoms.[42] The variability intrinsic to IBS symptoms, combined with which "snapshot" (i.e., relative symptom activity vs. quiescence) happens to be captured at the time of pre-post assessments clearly presents a challenge. Further, research trials often utilize questionnaire data for eligibility screening, whereas daily diaries may subsequently be used to maximize accuracy of participant's symptom data during a treatment trial, and those two methods or reporting intervals may not correspond. Though group-level bias may prove less problematic in trials evaluating pre-post treatment effects using the same measure, investigators should give thoughtful consideration to these methodological issues when designing clinical trials, as measurement choices may greatly impact results. By way of example, a recent randomized controlled trial of cognitive behavioral therapy (CBT) versus an intensive medical care (IMC) condition for treatment of pediatric functional abdominal pain illustrated the striking potential for method of assessment to affect conclusions. When assessing percentage of children improved or recovered directly post-treatment, questionnaire data reflected that 32% of children in the CBT condition were improved (30% in IMC), whereas diary data indicated that 67% were improved (47% in IMC).[43]

This study is limited in its relatively small sample size. Participants were also part of a treatment trial, participation in which could have affected patient experience or observation of symptoms, most likely improving the accuracy of their reporting by attending to and recording symptoms. It should also be noted that while we report percentages of exact agreement, in some circumstances less than perfect agreement may be "close enough."

In summary, our results suggest that for pain report, child recall questionnaire may be a reasonable approximation of child diary data when a short recall interval is used, though our conditions likely inflated agreement compared with the conditions routinely used in clinical practice and research, and the individual variability in children's accuracy of recall should be acknowledged. Further, relying on parent retrospective questionnaire report of child symptoms does not appear to be a suitable proxy, and recall of stool form by both parent and child may be more problematic than recall of pain. We believe these results combined with the limited existing literature and the unique features of IBS continue to advocate for use of prospective diary data to guide diagnosis and assessment whenever possible rather than relying on retrospective questionnaire.[24,44] The allure of efficiency is tempting, but relying on recalled pain and stool form may be misleading.[27,28] In addition to mitigating the problems already discussed, diaries also allow examination of sequence of events, which may be particularly important in IBS where pain-stool relations are a defining characteristic, the complexities of which may be more difficult for a reporter (especially a child or proxy reporter) to appreciate.[35]

The primary deterrent to diary use is respondent and provider burden and concerns about non-adherence, though the advent of electronic diaries offers methodological and logistical advantages that mitigate these problems at least to some extent.[45,46] Physicians and patients report positivity about the concept of using electronic diaries in clinical practice, but physicians admit they ultimately don't incorporate the data into treatment decision making due to time constraints or forgetting, and patients perceive that diary use does not impact care.[47] Development and evaluation of clinically feasible diaries is therefore a fruitful area for future research. Future research investigating factors affecting lack of agreement between recalled and recorded data also would be of conceptual interest. Qualities and variability of the specific symptom being assessed[12, 13,27,48,49], degree of symptom bother or interference[3,50], beliefs about typical symptom experience[4,24,25,51], or other individual characteristics[7] are examples with preliminary evidence of relevance. Finally, future research should incorporate how close agreement between methods must be to yield equivalent clinical judgments or empirical determinations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

Funding Statement

This research was supported by Grant Number R01 NR05337 and R01 NR013497 to RJS from the National Institutes of Health, the Daffy's Foundation, the USDA/ARS under Cooperative Agreement No. 6250-51000-043, and P30-DK56338 which funds the Texas Medical Center Digestive Disease Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work is a publication of the USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine and Texas Children's Hospital, Houston, TX. The contents of this publication do not necessarily reflect the views or policies of the USDA, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

REFERENCES

1. Irvine EJ, Whitehead WE, Chey WD, et al. Design of treatment trials for functional gastrointestinal disorders. Gastroenterology. 2006; 130:1538–1551. [PubMed: 16678567]
2. U.S. Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, & Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. 2009
3. Broderick JE, Stone AA, Calvanse P, et al. Recalled pain ratings: a complex and poorly defined task. The Journal of Pain. 2006; 7:142–149. [PubMed: 16459280]
4. Broderick JE, Schwartz JE, Vikingstad, et al. The accuracy of pain and fatigue items across different reporting periods. Pain. 2008; 139:146–157. [PubMed: 18455312]
5. Eich E, Reeves JL, Jaeger B, Graff-Radford S. Memory for pain: relation between past and present pain intensity. Pain. 1985; 23:375–379. [PubMed: 4088698]
6. Redelmeier DA, Kahneman D. Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. Pain. 1996; 66:3–8. [PubMed: 8857625]
7. Schneider S, Stone AA, Schwarts JE, et al. Peak and end effects in patients' daily recall of pain and fatigue: a within-subjects analysis. The Journal of Pain. 2011; 12:228–235. [PubMed: 20817615]

8. Salovey P, Smith AF, Turk DC, et al. The accuracy of memory for pain: not so bad most of the time. APS Journal. 1993; 2:184–191.
9. Weinland SR, Morris CB, Hu Y, et al. Characterization of episodes of irritable bowel syndrome using ecological momentary assessment. The American Journal of Gastroenterology. 2011; 106:1813–1820. [PubMed: 21647206]
10. Stone AA, Broderick JE, Katell AT, et al. Does the peak-end phenomenon observed in laboratory pain studies apply to real-world pain in rheumatoid arthritics? The Journal of Pain. 2000; 1:212–217. [PubMed: 14622620]
11. Jensen MP, Mardekian J, Lakshminarayanan M, et al. Validity of 24-h recall ratings of pain severity: Biasing effects of “Peak” and “End” pain. Pain. 2008; 137:422–427. [PubMed: 18035495]
12. Niere K, Jerak A. Measurement of headache frequency, intensity and duration: comparison of patient report by questionnaire and headache diary. Physiotherapy Research International. 2004; 9:149–156. [PubMed: 15790252]
13. Stone AA, Broderick JE, Schwartz JE. Validity of average, minimum, and maximum end-of-day recall assessments of pain and fatigue. Contemporary Clinical Trials. 2010; 31:483–490. [PubMed: 20620239]
14. Lewandowski AS, Palermo TM, Kirchner HL, et al. Comparing diary and retrospective reports of pain and activity restriction in children and adolescents with chronic pain conditions. Clin J Pain. 2009; 25:299–306. [PubMed: 19590478]
15. Peters ML, Sorbi MJ, Kruise DA, et al. Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. Pain. 2000; 84:181–192. [PubMed: 10666523]
16. Raselli C, Broderick JE. The association of depression and neuroticism with pain reports: A comparison of momentary and recalled pain assessment. Journal of Psychosomatic Research. 2007; 62:313–320. [PubMed: 17324682]
17. Stone AA, Broderick JE, Shiffman SS, et al. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. Pain. 2004; 107:61–69. [PubMed: 14715390]
18. van den Brink M, Bandell-Hoekstra ENG, Abu-Saad HH. The occurrence of recall bias in pediatric headache: a comparison of questionnaire and diary data. Headache. 2001; 41:11–20. [PubMed: 11168599]
19. Stone AA, Broderick JE, Schwartz JE, et al. Context effects in survey ratings of health, symptoms, and satisfaction. Medical Care. 2008; 46:662–667. [PubMed: 18580384]
20. Jamison RN, Raymond SA, Slawsby EA, et al. Pain assessment in patients with low back pain: comparison of weekly recall and momentary electronic data. The Journal of Pain. 2006; 7:192–199. [PubMed: 16516825]
21. Manning AP, Wyman JB, Heaton KW. How trustworthy are bowel histories? Comparison of recalled and recorded information. British Medical Journal. 1976; 2:213–214. [PubMed: 974496]
22. Heaton KW, Radvan J, Cripps H, et al. Defecation frequency and timing, and stool form in the general population: a prospective study. Gut. 1992; 33:818–824. [PubMed: 1624166]
23. Pamuk ON, Pamuk GE, Celik AF. Revalidaiton of description of constipation in terms of recall bias and visual scale analog questionnaire. Journal of Gastroenterology and Hepatology. 2003; 18:1417–1422. [PubMed: 14675272]
24. Ashraf W, Park F, Lof J, et al. An examination of the reliability of reported stool frequency in the diagnosis of idiopathic constipation. Am J Gastroenterol. 1996; 91(1):26–32. [PubMed: 8561138]
25. Bellini M, Bove A, Sormani MP, et al. The daily diary and the questionnaire are not equivalent for the evaluation of bowel habits. Digestive and Liver Disease. 2010; 42:99–102. [PubMed: 19473896]
26. Palsson OS, Baggish JS, Turner MJ, et al. IBS patients show frequent fluctuations between loose/watery and hard/lumpy stools: implications for treatment. Am J Gastroenterol. 2012; 107:286–295. [PubMed: 22068664]

27. Coletta M, Di Palma L, Tomba C, et al. Discrepancy between recalled and recorded bowel habits in irritable bowel syndrome. Alimentary Pharmacology & Therapeutics. 2010; 32:282–288. [PubMed: 20374227]
28. Lackner JM, J, Keefer L, et al. The accuracy of patient-reported measures for GI symptoms: a comparison of real time and retrospective reports. Journal of Neurogastroenterology & Motility. 2014; 26:1802–1811.
29. Lander J, Hodgins M, Fowler-Kerry S. Children's pain predictions and memories. Beh. Res. Ther. 1992; 30:117–124.
30. Lehmann HP, Bendebba M, DeAngelis C. The consistency of young children's assessment of remembered painful events. J. Dev. Behav. Pediatr. 1990; 11:128–134. [PubMed: 2365834]
31. Zonneveld LNL, McGrath PH, Reid GH, et al. Accuracy of children's pain memories. Pain. 1997; 71:297–302. [PubMed: 9231873]
32. Stinson JN, Jibb L, Lalloo C, et al. Comparison of average weekly pain using recalled paper and momentary assessment electronic diary reports in children with arthritis. The Clinical Journal of Pain. 2014 (in press).
33. Van der Plas TN, Benninga MA, Redekop WK, Taminiau JA, Buller HA. How accurate is the recall of bowel habits in children with defecation disorders. Eur J Pediatr. 1997; 156:178–181. [PubMed: 9083754]
34. Chogle A, Sztainberg M, Bass L, et al. Accuracy of pain recall in children. J Pediatr Gastroenterol Nutr. 2012; 55(3):288–291. [PubMed: 22314392]
35. Czyzewski DI, Lane MM, Weidler EM, Williams AE, Swank PR, Shulman RS. Interpretation of Rome III Criteria and Method of Assessment Affect the Irritable Bowel Syndrome Classification of Children. Alimentary Pharmacology and Therapeutics. 2011; 33(3):403–411. (2011). [PubMed: 21138454]
36. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol. 1997; 32:920–924. [PubMed: 9299672]
37. Cicchetti DV, Sparrow SS. Developing criteria for establishing the interrater reliability of specific items in a given inventory. American Journal of Mental Deficiency. 1981; 86:127–137. [PubMed: 7315877]
38. Cicchetti D, Bronen R, Spencer S, et al. Rating scales, scales of measurement, issues of reliability: resolving some critical issues for clinicians and researchers. Journal of Nervous and Mental Disease. 2006; 194:557–561. [PubMed: 16909062]
39. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; i:307–310. [PubMed: 2868172]
40. Rasquin A, Di Lorenzo C, Forbes D, et al. Childhood functional gastrointestinal disorders: Child/Adolescent. Gastroenterology. 2006; 130:1527–1537. [PubMed: 16678566]
41. Food and Drug Administration. Guidance for industry: Patient-reported outcome measures: Use in medical product development of support labeling claims. Federal Register. 2009. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.
42. Norquist JM, Girman C, Fehnel S, et al. Choice of recall period for patient-reported outcome (PRO) measures: criteria for consideration. Qual Life Res. 2012; 21:1013–1020. [PubMed: 21909804]
43. van der Veek SMC, Derkx BHF, Benninga MA, et al. Cognitive behavior therapy for pediatric functional abdominal pain: a randomized controlled trial. Pediatrics. 2013; 132:e1163–e1172. [PubMed: 24127467]
44. Naliboff BD. Choosing outcome variables: global assessment and diaries. Gastroenterology. 2004; 126:S129–S134. [PubMed: 14978649]
45. Palermo TM, Valenzuela D, Stork PP. A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability. Pain. 2004; 107(3):213–219. [PubMed: 14736583]
46. Stone AA, Broderick JE, Schwartz JE, et al. Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. Pain. 2003; 104:343–351. [PubMed: 12855344]

47. Marceau LD, Link CL, Smith LD, Carolan SJ, Jamison RN. In-clinic use of electronic pain diaries: barriers of implementation among pain physicians. J Pain Symptom Manage. 2010; 40(3):291–404.

48. Jensen MP, Turner LR, Turner JA, et al. The use of multiple-item scales for pain intensity measurement in chronic pain patients. Pain. 1996; 67:35–40. [PubMed: 8895229]

49. Stone AA, Schwartz JE, Broderick JE, et al. Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic. Pers Soc Psychol Bull. 2005; 31:1340–1346. [PubMed: 16143666]

50. Kenton K, Fitzgerald MP, Brubaker L. What is a clinician to do—believe the patient or her urinary diary? The Journal of Urology. 2006; 176:633–635. [PubMed: 16813908]

51. Conner TS, Barrett LF. Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. Psychosom Med. 2012; 74(4):327–337. [PubMed: 22582330]

Key Messages

- Recall of symptoms drives diagnostic decisions and evaluation of treatment response in pediatric functional gastrointestinal disorders, so accurately capturing patient and parent report of symptoms is important.
- This study evaluated correspondence between retrospective questionnaire (parent and child report) and two-week prospective diary data for 50 children and adolescents with Irritable Bowel Syndrome.
- Overall agreement between child questionnaire and diary was strong for pain and days without bowel movement, though with individual variation. Parent recall and child diary were less concordant, and agreement about diarrhea was poor for parent and child.
- Child questionnaire with short recall interval may reasonably approximate diary data for some individuals and some symptoms, whereas parent questionnaire does not appear a suitable proxy. These results combined with existing literature support use of diary data whenever possible.

Table 1

Means and Standard Deviations of Study Variables

	Child Diary	Child Questionnaire	Parent Questionnaire
Number of Pain Days	5.60 ± 4.38[a]	5.37 ± 4.06[a]	3.59 ± 4.40[b]
Maximum Pain Rating	4.82 ± 2.84[a]	4.68 ± 2.85[a]	3.80 ± 3.07[b]
Number of Days with No Bowel Movement	4.62 ± 3.74[a]	3.54 ± 3.37[b]	3.02 ± 3.53[b]
Number of Days with Diarrhea	0.70 ± 1.74[a]	0.89 ± 1.62[a]	0.73 ± 2.12[a]

Note. Differing superscripts reflect group differences at the $p < .05$ level per dependent samples *t*-tests.

Table 2

Agreement between Retrospective Questionnaire (Child and Parent-Report) and Diary Data

			Pain and Stooling Report		
	ICC	**Spearman's rho**	**Diary Exceeds Questionnaire**	**Questionnaire Exceeds Diary**	**Diary and Questionnaire Equal**
Number of Pain Days					
Child recall vs. diary	.83	.82*	40%	36%	24%
Number of Pain Days	.55	.53**	54%	22%	24%
Maximum Pain Rating					
Child recall vs. diary	.80	.76*	34%	26%	40%
Parent recall vs. diary	.42	.42*	46%	32%	22%
Number of Days with No BM					
Child recall vs. diary	.74	.77*	56%	20%	24%
Parent recall vs. diary	.74	.57*	58%	26%	16%
Number of Days with Diarrhea					
Child recall vs. diary	−.03	.12	14%	32%	54%
Parent recall vs. diary	.07	.16	20%	22%	58%

* $p<.01$

Table 3

Agreement of Diary and Child Questionnaire by Child Age Correlations between Child Age and Questionnaire vs. Diary Difference Scores (N = 50)

Difference Scores (Questionnaire – Diary)	Correlation with Age
Number of Pain Days	−0.11
Maximum Pain Rating	−0.01
Number of Days with No Bowel Movement	0.13
Number of Days with Diarrhea	0.19

All *ps* > .05

Table 4

Distribution of Discrepancy for Child-Reported Pain and Stooling Variables (N = 50)

	Number of Pain Days		
	Diary Exceeds Questionnaire	**Questionnaire = Diary**	**Questionnaire Exceeds Diary**
Maximum Pain Rating			
Diary Exceeds Questionnaire	18%	6%	10%
Questionnaire = Diary	12%	14%	14%
Questionnaire Exceeds Diary	10%	4%	12%
Number of Days with No Bowel Movement			
Diary Exceeds Questionnaire	22%	14%	20%
Questionnaire = Diary	8%	8%	8%
Questionnaire Exceeds Diary	10%	12%	8%
Number of Days with Diarrhea			
Diary Exceeds Questionnaire	6%	4%	4%
Questionnaire = Diary	22%	4%	16%
Questionnaire Exceeds Diary	12%	16%	16%